Journal of Biomolecular Techniques
16:167–169 © 2005 ABRF

PERSPECTIVES

CLIVE SLAUGHTER

A BRIGHT BUT DEMANDING FUTURE FOR CORE FACILITIES

Discussions among ABRF members during their February meeting in Savannah, Georgia, revealed concern about the future of resource facilities amidst the broadening scope of biomedical science as it evolves toward integrative "systems biology" [1–3]. Here I argue that, despite this anxiety, resource facilities will continue to play an expanding and increasingly essential role in biomedical research. Despite this reassuring outlook, however, I believe the character of the work undertaken by facilities will change, most particularly in increasing demands for intellectual involvement. ABRF should embrace this change with enthusiasm, and provide strong leadership to advance its progress and realize the opportunities it presents.

Among the most striking changes in biomedical science that have taken place over the last three decades is the widespread expansion in studies that involve collaboration between multiple groups of investigators. Such collaboration occurs when the work requires the participation of individuals with special expertise or access to special samples. An illustration of the extent of this change is provided by a sampling of papers in biology and biomedicine published by three journals of high visibility and prestige, *Nature, Science,* and *Cell,* over the last three decades (Table 1). Papers appearing in January 1975, 1985, 1995, and 2005 are classified as collaborative or noncollaborative according to the authors' institutional affiliations. The incidence of collaborative reports is seen to have risen from 25% to 74% during this period. Biology has indeed become a collaborative, multidisciplinary enterprise.

One facet of this profound change in biological research is the expansion of reliance upon core facilities. Centralized laboratories equipped with complex, expensive instruments and staffed by individuals skilled in acquisition and interpretation of the data derived from them are now commonplace in both academic and commercial institutions. Institutions establish core facilities because they recognize that advanced technologies must be provided to maintain or improve institutional competitiveness. They recognize that they can afford to provide these technologies only with centralization.

Systems biology has promoted this trend. Among the procedures essential to the systems approach now routinely conducted in core facilities are highly parallelized methods for genotyping,[4] RNA expression analysis,[5] protein expression analysis,[6] and metabonomics.[7] High throughput procedures for assigning functions to macromolecules and assessing macromolecular activity levels are also gaining more general currency. Examples include the use of RNA inhibition (RNAi) in arrays of cultured cells,[8] qualitative binding

TABLE 1

Changing Proportion of Papers in Biology and Biomedical Science Representing Collaborative Work over the Last Three Decades

Year[a]	Collaborative Studies[b]	Noncollaborative Studies[c]	Total
1975	40 (25%)	119 (75%)	159
1985	67 (49%)	69 (51%)	136
1995	60 (61%)	38 (39%)	98
2005	70 (74%)	25 (26%)	95

[a]Figures based on sampling of the January issues of Nature, Science, and Cell (January and February issues of Cell are combined for 1975 and 1985)

[b]Papers authored by individuals belonging to different departments or institutions. N.B. This underestimates the incidence of collaborations to the extent that collaborating principal investigators with institutional affiliations in common are not counted.

[c]Papers authored by individuals with shared affiliations.

assays in protein microarray platforms,[9] and quantitative binding measurements made by surface plasmon resonance.[10] It is indeed through the implementation of such techniques that systems biology has become established as a communal aspiration of biomedical science.

I contend that the trend toward increasing reliance upon core facilities will persist, for they will continue to be called upon to implement further novel experimental procedures for systems biology. This task will involve streamlining the techniques, reducing turn-around times, teaching students how to use the new methods, and working with manufacturers to reduce costs.

When these efforts are successful, some technologies may be amenable to translation to commercial and clinical laboratories. For example, considerable oligonucleotide and peptide synthesis is presently performed in large commercial houses. Such divestment, I believe, frees core laboratories to fulfill the role for which they are best suited—the implementation of new and exciting vehicles for biological discovery—and represents a benefit rather than a detriment to core facilities.

The challenge for core facilities, then, is to find ways to function effectively in an environment where the technologies they are being called upon to deploy are crucial for driving biomedical research forward, yet are still to some extent immature. I believe that to be successful in this environment requires first a willingness to become proactively involved in projects at the highest level of intellectual input. The background and goals of every project must be understood in detail in order to apply the most appropriate techniques. Data must be evaluated not only at the technical level. The strength of the biological conclusions that can be drawn must also be assessed. For example, the interpretation of microarray expression data demands attention both to data quality and to statistical methods for data interpretation. Similarly, the elucidation of protein interactions by co-immunoprecipitation demands not only the ability to interpret the mass spectral data both qualitatively and quantitatively, but also careful attention to precipitation conditions and formulation of appropriate controls both negative and positive.

Facilities' success also depends upon planning for the future based upon a thorough understanding of the nature of the biological questions being addressed and in-depth familiarity with the latest methodological literature. Core facilities must accept responsibility for identifying the most effective methods of accomplishing research goals. Simply continuing to perform procedures derived from today's technology without anticipating and working to implement tomorrow's methods will consign a facility to obsolescence. Careful scientific judgment is required to make such assessments, and correctly anticipating demands for new capabilities is vital to scientific productivity. Oftentimes, more than one platform is available for acquiring the information that is needed, and it falls to facility personnel to make or guide choices between alternative methods. For example, it has only recently become clear that genotyping and loss of heterozygosity are likely henceforth to be accomplished most effectively with DNA microarrays in a high throughput context,[4] but, until recently, capillary electrophoresis and mass spectrometry were also vying for attention.[11] Aside from scientific productivity, the decision to invest in a platform can have important financial consequences. In the future, what will be the best methods for identifying and assessing the function of small RNA species? How will the dynamic range problem in proteomics be definitively solved? Such questions should be the subject of active inquiry by core facilities right now.

Success further depends upon recruiting staff with a high level of scientific sophistication, and hence garnering the salaries and benefits commensurate with such skills. Core facility management requires additional expertise and possibly special training in project management techniques. Furthermore, operating as part of a multidisciplinary team often requires considerable skill in communicating information to collaborators, and not infrequently requires considerable diplomacy. It is greatly advantageous to structure core labs in ways that promote communication between individuals working on interdependent aspects of a single project. Multidisciplinary cores tend to promote such communication.

Another key requirement for deploying many contemporary techniques is computational support, including access to suitable computers and software, help with database management, and availability of statistical or bioinformatic expertise. Adequate resources of these kinds may require expenditures equal to or even greater than those made in deploying the analytical methods themselves. Among the reasons are the exceedingly heavy demands on processor and communications speed, memory, and storage space, and the specialized nature of the skills to make best use of the hardware involved.

Clearly, institutional resources are needed to create core laboratories. Once a decision to invest in a technology has been made, the arguments for centralizing its availability in a core facility are frequently compelling. These arguments include the ability to regulate access in a way that is equitable on an institution-wide basis, and the control of costs through shared infrastructure, including computational support. Although the cost may be high, the benefits of creating produc-

tive core laboratories are correspondingly great, and include enhanced productivity and visibility of research, and competitive advantage in recruiting excellent faculty.

Success in this evolving environment does not entail elimination of small core facilities. At the February meeting of ABRF in Savannah, voice was given to fears that small cores are liable to be "squeezed out" in the era of systems biology. However, I see compelling reasons to believe that small cores will continue to be formed to pioneer and deploy new, special methods, and will grow as their work becomes recognized. Some may diversify their services as growth takes place. The existing large multidisciplinary cores all arose in this way. Others will not diversify in this way, but will nonetheless continue to provide crucial services.

A primary goal of ABRF, I believe, should be to assist its member cores in meeting these challenges. Among the specific actions that would provide most benefit at the present time are the following:

- Provision of listings of nationally and internationally recognized scientists as consultants in relevant fields who are willing to participate in scientific advisory boards and review committees for member cores;
- Coordination of research aimed at comparing alternative approaches to solving contemporary problems—a goal that Research Committees should embrace as their primary concern;
- Development of training programs in project management and contemporary techniques for statistical analysis and database management;
- Development of policy recommendations for research funding to be promulgated through FASEB; and
- Identification of laboratories that do not self-identify as cores but which can strengthen ABRF in meeting the challenges of its members in the present era of change. Possible examples are to be found in the fields of imaging, small molecule mass spectrometry, small molecule NMR spectroscopy, and clinical chemistry.

We are at a juncture when the opportunities for core laboratories to expand and to extend the contributions they make to biomedical science are unprecedented, but cores must adapt and strengthen their capabilities to meet the challenges those opportunities present.

REFERENCES

1. Westerhoff HV, Palsson BO. The evolution of molecular biology into systems biology. *Nat Biotechnol* 22; 2004:1249–1252.
2. Butcher EC, Berg EL, Kunkel EJ. Systems biology in drug discovery. *Nat Biotechnol* 22;2004:1253–1259.
3. Weston AD, Hood L. Systems biology, proteomics, and the future of health care: Toward predictive, preventative, and personalized medicine. *J Proteome Res* 3;2004: 179–196.
4. Broman KW, Feingold E. SNPs made routine. *Nat Methods* 1;2004:104–105.
5. Stoughton RB. Applications of DNA microarrays in biology. *Annu Rev Biochem* 74;2005:53–82.
6. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 422;2003:198–207.
7. Lindon JC, Holmes E, Nicholson JK. So what's the deal with metabonomics? *Anal Chem* 75;2003:384A–391A.
8. Wheeler DB, Bailey SN, Guertin DA, Carpenter AE, Higgins CO, Sabatini DM. RNAi living-cell microarrays for loss-of-function screens in *Drosophila melanogaster* cells. *Nat Methods* 1;2004:127–132.
9. Phizicky E, Bastiaens PIH, Zhu H, Snyder M, Fields S. Protein analysis on a proteomic scale. *Nature* 422;2003: 208–214.
10. Day YSN, Baird CL, Rich RL, Myszka DG. Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods. *Protein Sci* 11;2002: 1017–1025.
11. Syvänen A-C. Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat Rev Genet* 2;2001: 930–942.

Journal of Biomolecular Techniques
16:170 © 2005 ABRF

RESEARCH GROUP/ COMMITTEE REPORTS

NUCLEIC ACIDS RESEARCH GROUP

In January 2005, the NARG finished analysis of the data returned from the participants of the 2004/2005 research project., *Validation of Your Reverse Transcription Real-Time PCR Technique*. Results were presented in talks by Scottie Adams and Brian Holloway at the research group presentation, "A Comparison of Real-time PCR Techniques, Chemistries, and Hardware in Laboratories Utilizing the Same Assay," as well as in a poster, at ABRF 2005 in Savannah GA. The research group presentation was followed by a discussion on real-time PCR problems, led by Greg Shipley and Stephen Bustin. Deborah Grove chaired a round table discussion, Lawrence Wangh of Brandeis spoke on "LATE PCR and Allied Technologies," and Reinhold Mueller of Stratagene discussed "Process Considerations of Sample Preparation to Normalized Gene Expression Profiling." Greg Shipley chaired a tutorial entitled "Utilizing qPCR to Validate Microarray Results" with invited speakers Tim Hunter of the University of Vermont and Andrew Brooks of the University of Rochester. Titles of their talks were "Validation of Microarray Data via Quantitative Real-Time PCR" and "High-Throughput Real-Time PCR Approaches for Microarray Assessment and Target Validation."

The group elected Deborah Grove of the Pennsylvania State University as new chair and welcomed Tim Hunter as a new member. The NARG thanks Brian Holloway of the Centers for Disease Control for his many years of service on the committee and transferred his status to ad hoc. Also, the group thanks Stephen Scaringe, ad hoc member, who has completed his membership on the committee.

The group is discussing ideas for the 2005/2006 NARG research project and will be pursuing possibilities of joining in projects suggested by FARG and DSRG as well as a project with MARG.

PROTEOMICS RESEARCH GROUP

The PRG welcomes its newest members: Ewa Witkowska from the University of California, San Francisco and Nathan Yates from Merck Research Laboratories. We also thank outgoing member Karen West and outgoing chair Tom Neubert for their much appreciated service during the last two and three years, respectively. Chris Turck has generously agreed to stay on the PRG for a fourth year and will take over as Chair for the coming year.

This year's PRG study to evaluate participants' ability to determine the sequences of peptides *de novo* has been completed, and the results were presented at the ABRF '05 meeting in Savannah in a talk by Brett Phinney and on a poster by the group. A mixture containing 3–6 pmol each of five synthetic peptides, the sequences of which were not present in any public database, was distributed to 106 requesting laboratories. Data were returned from 40 laboratories resulting in a 37% return rate. As is often the case, the study proved to be more difficult than anticipated by the PRG; de novo sequencing of proteins remains a substantial challenge. Results of the study as well as current information about the activities of the PRG can be found on the PRG web site at http://www.abrf.org/index.cfm/group.show/Proteomics.34.htm.

Journal of Biomolecular Techniques
16:171–177 © 2005 ABRF

ARTICLE WATCH

This column highlights recently published articles that are of interest to the readership of this publication. We encourage ABRF members to forward information on articles they feel are important and useful to Clive Slaughter, Hartwell Center, St. Jude Children's Research Hospital, 332 North Lauderdale St., Memphis, TN 38105-2794; Tel: (901) 495-4844; Fax: (901) 495-2945; Email: Clive.Slaughter@stjude.org; or to any member of the Editorial Board. Article summaries reflect the reviewers' opinions and not necessarily those of the Association.

AMINO ACID ANALYSIS AND PROTEIN SEQUENCING

Alterman MA, Gogichayeva NV, Kornilayev BA. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry-based amino acid analysis. *Analytical Biochemistry* 335;2004:184–191.

Quantitation of amino acids from protein acid hydrolysates is demonstrated using MALDI-TOF mass spectrometry with α-cyano-4-hydroxycinnamic acid as matrix. All the standard protein amino acids that are measured using conventional chromatographic separation procedures are included, except that leucine and isoleucine are not discriminated. No ion suppression effects are observed with the methods described. Methyltyrosine is used as internal standard. Linear responses are documented between 20 and 300 μM concentration with correlation coefficients between 0.983 and 0.999. Limits of quantitation are between 0.03 μM and 3.7 μM. The main advantages of the method are that it requires no derivatization or chromatographic separation of the amino acids. Data acquisition is therefore exceedingly fast.

Samyn B, Sergeant K, Castanheira P, Faro C, van Beeumen J. A new method for C-terminal sequence analysis in the proteomic era. *Nature Methods* 2;2005:193–200.

Proteins in solution, in gels, or on polyvinyl difluoride blots are digested with cyanogen bromide under conditions that maximize the abundance of homoserine lactone relative to homoserine at the resulting peptide C-termini. The unfractionated peptide mixture is then digested with carboxypeptideases Y and P. Peptides with C-terminal homoserine lactone are refractory to digestion. The peptide derived from the C-terminus of the intact protein, however, is cleaved to give a ladder of signals revealed by MALDI-TOF mass spectrometry from which the C-terminal sequence can be deduced. The technique is convenient in circumstances where mass measurements of carboxypeptidase digests of the intact protein are impractical to perform with sufficient sensitivity and accuracy.

DNA CHARACTERIZATION AND GENOTYPING

Matsuzaki H, *et al.* Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nature Methods* 1;2004:109–111.

Genotyping of 116,204 human single nucleotide polymorphisms (SNPs) using Affymetrix oligonucleotide arrays is documented. Genomic DNA is digested using *Xba*I and *Hin*dIII separately, in parallel. Restriction fragments are ligated to adaptors and amplified with Platinum *Pfx* polymerase from Invitrogen to produce fragments of 250–2000 bp. Each SNP is represented on the arrays by 40 probes—10 groups of 4 representing perfect match and mismatch pairs for both alleles. Genotyping is accomplished by an algorithm that calculates the log-likelihood of the possible genotypes (homozygotes AA and BB, heterozygote AB, and null) based on the observed hybridization intensities to the groups of 4. Suitable confidence limits for acceptance are imposed. The average distance

between markers is 23.6 kb, and 92% of the genome is estimated to be within 100 kb of a marker. The methodology provides a broadly applicable vehicle for genome-wide association studies.

Murray SS, Oliphant A, Shen R, McBride C, Steeke RJ, Shannon SG, Rubano T, Kermani G, Fan J-B, Chee MS, Hansen MST. A highly informative SNP linkage panel for human genetic studies. *Nature Methods* 1;2004:113–117.

A panel of 4763 SNP markers is presented for genetic linkage studies within families. Genotyping is performed with arrays of beads using the Illumina BeadArray technology. Each bead type has a specific oligonucleotide probe sequence attached. The arrays are randomly assembled collections of beads, such that each bead type is represented an average of 30 times per array. Each bead is located in a microwell at the end of an optical fiber. The fibers are bundled, and bundles are arranged in a 96-well format with a different array in each well for convenience in processing multiple samples in parallel. This system is being used to generate half of the genotyping data in the International HapMap Project. The SNP panel is used in the present work to type 518 individuals in 28 large families and hence to construct a genetic map of the markers. The mean genetic map distance between markers is 1.5 cM. The system provides a broadly applicable, high-throughput platform for linkage studies of disease markers.

CARBOHYDRATES, GLYCOLIPIDS, AND GLYCOPROTEINS

Larsen MR, Hojrup P, Roepstorff P. Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Molecular and Cellular Proteomics* 4;2005:107–119.

Gel bands of N-linked glycoprotein are subjected to a two-stage digestion procedure. First, trypsin is employed and an aliquot of the digest used for mass spectrometric protein identification. The remainder of the digest is further treated with the nonspecific proteinase K to yield small peptides. Unglycosylated peptides are removed upon passage through a microcolumn consisting of a GELoader tip (Eppendorf) packed with Poros 2. Glycosylated peptides are not retained, but are trapped on a second GELoader tip packed with graphite powder, washed to remove low molecular weight contaminants, and eluted with acetonitrile. Tandem mass spectrometry then provides amino acid sequence and partial glycan structure. Using this strategy on 8 pmol of ovalbumin applied to a gel, all 13 of the previously known glycan chains were identified, plus three additional ones.

Zhang H, Yi EC, Li X-j, Mallick P, Kelly-Spratt KS, Masselon CD, Camp DG II, Smith RD, Kemp CJ, Aebersold R. High throughput quantitative analysis of serum proteins using glycopeptide capture and liquid chromatography mass spectrometry. *Molecular and Cellular Proteomics* 4;2005:144–155.

Quantitative serum proteomics based on LC/MS analysis of tryptic peptides from serum proteins is compromised by the extreme complexity of the peptide mixtures produced by trypsin digestion. The present work offers a method of simplifying the mixture. Glycoproteins are immobilized by periodate oxidation of hydroxyl groups on their sugar side chains to aldehydes, then covalently binding them to hydrazide beads. The immobilized proteins are then digested with trypsin, and unbound peptides are washed away. The glycopeptides remaining immobilized are then released by digestion with peptide-N-glycosidase F, and are employed for quantitative LC/MS analysis. By restricting attention to those peptides that were previously glycosylated, the complexity of the peptide mixture is reduced, and the sensitivity and throughput of analyses are increased.

MACROMOLECULAR SYNTHESIS

Lausted C, Dahl T, Warren C, King K, Smith K, Johnson M, Saleem R, Aitchison J, Hood L, Lasky SR. POSaM: A fast, flexible, open-source, inkjet oligoncleotide synthesizer and microarrayer. *Genome Biology* 5;2004:R58.

Noting the spur given to the microarray field by the early release of the Stanford design for a pin-spotting arrayer, and drawing attention to the advantages of synthesizing oligonucleotides in situ on chips to create custom arrays, the authors seek to remedy the lack of ready access to instrumentation for performing in situ, custom oligonucleotide synthesis in academic laboratories. A piezoelectric oligonucleotide synthesizer and arrayer is described. It uses a low-cost print head, high-quality motion controllers, and standard phosphoramidite chemistry, and rapidly produces arrays of 9800 features. The construction can be undertaken by most well-equipped molecular biology laboratories with modest organic chemistry and engineering expertise.

Cline DJ, Thorpe C, Schneider JP. General method for facile intramolecular disulfide formation in synthetic peptides. *Analytical Biochemistry* 335;2004:168–170.

4,4′-Dithiodipyridine is shown greatly to facilitate oxidation of peptides to form intramolecular disulfide bonds. The reaction takes place under the acidic conditions and in the presence of high concentrations of organic solvents that typify the conditions under which peptides are commonly purified by reverse-phase chromatography. This method for disulfide bond formation can therefore rapidly be performed following purification with a minimum of intervening manipulation.

METABONOMICS

Jackson SN, Wang H-YJ, Woods AS, Ugarov M, Egan T, Schultz JA. Direct tissue analysis of phospholipids in rat brain using MALDI-TOFMS and MALDI-ion mobility-TOFMS. *Journal of the American Society of Mass Spectrometry* 16;2005:133–138.

Frozen brain tissue is sectioned on a cryostat. Sections are placed on a target plate and spots of matrix solution—optimally, 6-aza-2-thiothymine or 2,6-dihydroxyacetophenone—are applied. Ions of three lipid classes, phosphatidylcholines, phosphatidylethanolamines, and sphingomyelin, are recorded in the resulting MALDI mass spectra. PC 32:0, PC 34:1, and SM 18:0 predominate. The identities of the phospholipid ions are confirmed in ion mobility studies. The method is amenable to use in tissue imaging, and to the detection of lipophilic drugs.

MASS SPECTROMETRY

Kong XL, Huang LCL, Hsu C-M, Chen W-H, Han C-C, Chang H-C. High-affinity capture of proteins by diamond nanoparticles for mass spectroscopic analysis. *Analytical Chemistry* 77;2005:259–265.

Proteins are concentrated from dilute solution by virtue of their ability to interact with diamond nanoparticles. Diamond powder with nominal particle size of 100 nm is treated with strong oxidative acid to yield a particle surface that is carboxylated and oxidized. The particles are hydrophilic and stable to storage in aqueous suspension. Protein complexed to diamond is sedimented by centrifugation, and is amenable to analysis by MALDI without prior removal of the particles.

Cargile BJ, Bundy JL, Stephenson JL. Jr. Potential for false positive identifications from large databases through tandem mass spectrometry. *Journal of Proteome Research* 3;2004:1082–1085.

There is a general and growing concern about the assumption that protein identifications made in shotgun proteomics are unquestionably correct. The authors here create a protein sequence database of a mythical creature (Medusa) consisting of 40,000 randomly generated protein sequences. The sizes of the proteins and the frequencies of the amino acids are made to emulate the human proteome. This database is used for "identification" of proteins from rat testis based on searches of MS/MS spectral data derived from tryptic peptides. Using scoring cut-offs of moderate stringency, 1400 proteins are "identified" with SEQUEST and 500 with MASCOT. Considering only those proteins "identified" on the basis of two or more peptides, 30 proteins are identified with SEQUEST and 2 proteins with MASCOT. Not all the matches can be eliminated by manual spectral validation. Such spurious results are to be anticipated because large databases, such as those available for eukaryotes, contain enough peptides for random matches to be likely to occur with a high degree of significance. Of equal concern is the danger that correct identifications are missed through attempts to eliminate false identifications by elevating scoring cut-offs. The authors advocate introducing additional matching criteria, such as peptide isoelectric point, to eliminate false-positives, and hence to allow scoring cut-offs to be relaxed without risking a large increase in false identifications. Peptides may be selected by isoelectric focusing, and the measured pI values compared with those computed for sequences matched on the basis of mass spectrometry.

PROTEOMICS

Leichert LI, Jakob U. Protein thiol modifications visualized in vivo. *PloS Biology* 2;2004:e333.

Sethuraman M, McComb ME, Huang H, Huang S, Heibeck T, Costello CE, Cohen RA. Isotope-coded affinity tag (ICAT) approach to redox proteomics: Identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *Journal of Proteome Research* 3;2004:1228–1233.

Both these papers present methodology for identifying proteins whose activity is regulated by redox

potential through reversible modification of reactive cysteine residues. Disulfide bond formation, nitrosylation, glutathionylation, or sulfenic acid formation cause conformational changes that lead to protein activation or inactivation as part of the cellular response to oxidative stress. Leichert and Jakob lyse cells in trichloroacetic acid to quench thiol exchange. They then alkylate cysteines present in reduced state using iodoacetamide. Cysteines present in an oxidized state are then reduced with dithiothreitol and reacted with ^{14}C-labeled iodoacetamide. The radioactivity of proteins separated by 2-dimensional gel electrophoresis then provides an indication of their oxidation status. Comparison between control cells and cells subjected to oxidative stress identifies proteins putatively subject to redox regulation. Sethuraman et al. measure susceptibility of protein thiols to oxidation in vitro. Cysteines are labeled with the Applied Biosystems' cleavable ICAT reagent before and after treatment of tissue homogenates with hydrogen peroxide. Only peptides containing reduced cysteine react with the reagent, so cysteines susceptible to oxidation are recognized by depletion of the corresponding ICAT-labeled peptides.

Ferguson RE, Carroll HP, Harris A, Maher ER, Selby PJ, Banks RE. Housekeeping proteins: A preliminary study illustrating some limitations as useful references in protein expression studies. *Proteomics* 5;2005:566–571.

Variation in the expression levels of proteins commonly used as internal controls in protein expression studies is examined. Glyceraldehyde-3-phosphate dehydrogenase, β-actin, β-tubulin, and class I β-tubulin are quantitated by Western blotting as a function of total protein level in a series of renal cancer cell lines, in matched pairs of renal tumors and normal kidney cells, and in nine different human tissues. Each of the markers tested varies significantly in at least one biological context, often for reasons concerned with the various cellular functions they fulfill. It is therefore recommended that the choice of housekeeping proteins for internal standards is made carefully in relation to the cell and tissue types, experimental conditions, and disease states under consideration.

Dunkley TPJ, Watson R, Griffin JL, Dupree P, Lilley KS. Localization of organelle proteins by isotope tagging (LOPIT). *Molecular and Cellular Proteomics* 3;2004: 1128–1134.

Unambiguous assignment of the subcellular localization of proteins has long been problematic because pure preparations of organelles are exceedingly difficult to make. This paper presents a method for assigning subcellular localization that is suitable for use with organelles only partially separated by centrifugation through self-generating density gradients. Proteins sharing a common subcellular distribution are expected to exhibit a similar distribution in the gradient. Protein distributions are measured in pairwise comparisons of gradient fractions using the Applied Biosystems' cleavable ICAT reagent with mass spectrometric quantitation. Multivariate data analysis is then used to match these distributions to the distributions of known organelle-specific markers to assign subcellular localization.

Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha X-M, Polakiewicz RD, Comb MJ. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nature Biotechnology* 23;2005:94–101.

Gembitsky DS, Lawlor K, Jacovina A, Yaneva M, Tempst P. A prototype antibody microarray platform to monitor changes in protein tyrosine phosphorylation. *Molecular and Cellular Proteomics* 3;2004:1102–1118.

These papers present different strategies for approaching the same problem—the global profiling of proteins that are subject to tyrosine phosphorylation. Rush et al. digest cell extracts with proteases such as trypsin, then isolate peptides containing phosphotyrosine by immunoaffinity chromatography using a pTyr-specific monoclonal antibody. The peptides are then identified by mass spectrometry. In this way, the pool of species containing a modification of relatively low abundance is enriched, allowing identification of many previously unknown modification targets. Gembitsky et al. detect changes in phosphorylation state of known tyrosine-phosphorylated proteins using an antibody array method. Antibodies against specified phosphoproteins are arrayed as capture antibodies. The array is then incubated with a whole-cell or tissue extract, and probed with a fluorescently labeled phosphotyrosine-specific monoclonal antibody. This method is capable of high sensitivity and throughput, and is amenable to multiplexing.

Zappacosta F, Annan RS. N-terminal isotope tagging strategy for quantitative proteomics: Results-driven analysis of protein abundance changes. *Analytical Chemistry* 76;2004:6618–6627.

An isotope tagging method is described that labels every peptide from a protein in a sequence-independent manner, and is also suitable for peptides containing posttranslational modifications. Peptides in a protein digest are subjected to acylation with either d_0- or d_5-propionic anhydride. Lysine side chains are blocked prior to the reaction, so labeling is restricted to the peptide N-terminus. Protein abundance differences are detected by mass spectrometric analysis of mixtures of isotopically distinguished pep-

tides, and may be identified by data-dependent acquisition of MS/MS spectra.

MICROARRAYS

Choe SE, Boutros M, Michelson AM, Church GM, Halfon MS. Preferred analysis methods for Affymetrix GeneChips revealed by a wholly defined control dataset. *Genome Biology* 6;2005:R16.

Choices between alternative methods for analysis of gene expression data to optimize sensitivity while minimizing false discovery rates can be made only by investigating differences in the performance of analysis methods in cases where the RNA abundance levels are already known. The present paper evaluates analysis options with wholly defined RNA mixtures. A mixture of 2551 defined RNA species provides a constant background, while 100–200 RNAs are spiked in at levels representing fold-changes from 1.2 to 4.0. Accurate estimates of false-positive and false-negative rates can thus be made at each fold-change level, and nonspecific signal strength can be measured using probe sets corresponding to RNAs that are genuinely absent from the sample. All normalization methods perform similarly. However, substantial improvements to false-negative and false-positive discovery rates are obtained by subtracting nonspecific signal from the perfect match probe intensities, performing an intensity-dependent normalization at the probe set level, and incorporating an intensity-dependent standard deviation in the test statistic.

Mukherjee S, Berger MF, Jona G, Wang XS, Muzzey D, Snyder M, Young RA, Bulyk ML. Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nature Genetics* 36;2004:1331–1339.

Liu X, Noll DM, Lieb JD, Clarke ND. DIP-chip: Rapid and accurate determination of DNA-binding specificity. *Genome Research* 15;2005:421–427.

These papers introduce the use of DNA microarrays for high throughput identification of transcription factor binding sites. The method of Mukherjee et al. measures the direct binding of transcription factors to DNA microarrays. Epitope-tagged transcription factors are applied to a spotted whole-genome intergenic array, and then detected using fluorescently labeled, tag-specific antibodies. The data are normalized relative to the amount of double-stranded DNA in each microarray feature by measuring binding of the dye SybrGreen I to a duplicate array. Three yeast transcription factors are studied, and, in addition to the detection of known binding sequences, new targets are identified in each case. Many are upstream of previously uncharacterized open reading frames, and many are evolutionarily conserved. Liu et al. isolate protein-DNA complexes by immunoprecipitation from an in vitro mixture of genomic DNA and the pure transcription factor. The DNA fragments are then identified using a whole-genome microarray. The method is validated using a yeast transcription factor of known specificity.

Cleary MD, Meiering CD, Jan E, Guymon R, Boothroyd JC. Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. *Nature Biotechnology* 23;2005: 232–237.

Steady-state levels of macromolecules are governed by their combined rates of synthesis and degradation. Furthermore, the speed with which a cell can respond to a change in circumstances, determined, for example, by the rate at which it can switch between different steady-state levels of a regulatory macromolecule, also depends on both synthesis and degradation rates of that macromolecule. The present paper describes a method for measuring the rates of synthesis and degradation of mRNAs that is compatible with standard microarray-based methods for determining mRNA abundance. The method utilizes an enzyme in the protozoan parasite, *Toxoplasma gondii*, UPRT, to add a phosphoribosyl group to thiouridine to form thio-UMP, which is then available for incorporation into RNA. UPRT is shown to be amenable to expression and function in human cells. Thiouracil is rapidly taken up by cells, and is also rapidly chased out by uracil, with no discernable effect on gene expression levels. Thio-RNA is isolated by labeling with biotin and then performing avidin affinity chromatography. Rates of synthesis are measured by applying the thio-RNA made during thiouracil pulse to a standard expression microarray, and rates of degradation are measured from mRNA made during the subsequent uracil chase.

Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, Gerstein M, Snyder M. Global identification of human transcribed sequences with genome tiling arrays. *Science* 306;2004:2242–2246.

This paper provides a draft expression map for the entire human genome. Using maskless photolithography, 134 high-density oligonucleotide arrays are made representing 1.5 Gb of nonrepetitive genomic DNA. The probes are 36-mers and are synthesized with a feature density of 390,000 probes per array, and are positioned on average every 46 residues. The total number

of probes is 51,874,388. Transcribed sequences are located by hybridizing the arrays to fluorescently labeled cDNA reverse-transcribed from poly(A^+) RNA from pooled liver tissue. A total of 10,595 novel transcripts are identified, many of which are believed to be functional because of their homology with known mouse proteins. Many are located in regions distal to known genes. Some encode proteins of 300 amino acids or more. The remaining transcripts presumably encode small proteins, untranslated exons, or RNA species with presently unknown function.

FUNCTIONAL GENOMICS AND PROTEOMICS

Kim D-H, Behlke MA, Rose SD, Chang M-S, Choi S, Rossi JJ. Synthetic dsRNA dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnology* 23;2005: 222–226.

Siolas D, Lerner C, Burchard J, Ge W, Linsey PS, Paddison PJ, Hannon GJ, Cleary MA. Synthetic shRNAs as potent RNAi triggers. *Nature Biotechnology* 23;2005:227–231.

These two papers report improvements in the design of RNAs to provide major enhancements in effectiveness as agents for RNA interference (RNAi). Kim et al. show that 25–30-mer duplexes can be up to 10-fold more potent than the corresponding 21-mer duplexes normally used, and that some sites not susceptible to silencing by 21-mers are effectively targeted by 27-mers, with silencing lasting as long as 10 days. They did not observe induction of interferon response or activation of protein kinase R. Siolas et al. show that short hairpin RNAs with 29-bp stems and 2-bp 3′ overhangs are also more effective than conventional reagents. These features are believed to work by making the RNAs better substrates for cleavage by the processing enzyme, Dicer.

Wheeler DB, Bailey SN, Guertin DA, Carpenter AE, Higgins CO, Sabatini DM. RNAi living-cell microarrays for loss-of-function screens in *Drosphila melanogaster* cells. *Nature Methods* 1;2004:127–132.

Earlier array-based methods for screening living cells for RNAi-based loss of function have employed mammalian cells transfected with vectors expressing 21–23-mer short interfering RNAs or short hairpin RNAs to avoid sequence-independent activation of interferon reponse. In the present work, slide chemistries are modified to permit the growth of *Drosophila* cells, in which RNAi can, instead, be induced by long, double-stranded (ds) RNAs, with very high efficiency and specificity. Genome-wide collections of dsRNAs are available for *Drosophila*, and a prototype array of 384 features is used in the present study. In the case of *Drosophila* cells, two distinct dsRNAs can be used to silence two genes simultaneously, allowing the method to be used to identify interactions of genetic suppressors, enhancers, and synthetic lethals.

BIOINFORMATICS

Shih JH, Michalowska AM, Dobbin K, Ye Y, Qui TH, Green JE. Effects of pooling mRNA in microarray class comparisons. *Bioinformatics* 20;2004:3318–3325.

Pooling RNA from different samples for microarray-based class comparisons is done either when insufficient RNA from each individual is available for testing on its own array, or when the number of arrays is reduced to minimize costs. This paper assesses the consequences of pooling for the power of an experiment to detect differential gene expression between classes of individuals, taking into account the magnitude of experimental variation compared with the variation between individuals. To offset the loss of degrees of freedom due to pooling, multiple pools from different individuals must be used to achieve the same statistical power. The smaller the number of independent pools employed, the larger the number of individual samples the pools must contain to achieve comparable power. The savings of costs achieved by pooling may be outweighed by the added numbers of individuals required, depending on the relative costs of samples and chips. Formulae relating these variables are supplied to assist in experimental design.

Chang J, Van Remmen H, Ward WF, Regnier FE, Richardson A, Cornell J. Processing of data generated by 2-dimensional gel electrophoresis for statistical analysis: Missing data, normalization, and statistics. *Journal of Proteome Research* 3;2004:1210–1218.

In protein expression studies performed by 2D gel electrophoresis, it is common to observe spots that cannot be matched across all gels in a series for the purpose of quantitative comparison of spot staining intensity. Unmatched spots may arise through decreased quantity of the affected proteins, change in migration due to post-translational modification, experimental variations such as insufficient resolution, or failure to detect a spot because it falls below the intensity detec-

tion threshold. The resultant "missing data" present the problem of how to include unmatched spots in quantitative assessments of protein expression change. The present work suggests using a method involving K-nearest-neighbors–based imputation, and describes conditions under which the method should be invoked. Alternative methods for normalization and for statistical testing are also evaluated.

Journal of Biomolecular Techniques
16:178 © 2005 ABRF

NEWS & EVENTS

MESSAGE FROM THE EXECUTIVE BOARD

Fellow members of the ABRF, first I wish to introduce myself, Jay Fox, as the current president of the ABRF. I assumed the position at the end of the national meeting in Savannah in February. I have had a long, useful, and enjoyable association with the ABRF and it is an honor to serve as the president of the association. I know that you all will join me in thanking Kristine Swiderek for her outstanding contributions as president of the organization for the last two years. She has guided the Executive Board and the ABRF with insight and a steady purpose that will continue to benefit the future of our organization.

I wish to report that the Association is in a strong position with regard to its membership, fiscal position, and activity level among the committees and research groups. Over the past year, the Executive Board has launched the new Computational Biology Research Group, and in concert with the current Proteomics Research Group, expanded that area to include a Proteomics Standards Research Group and Quantitative Proteomics Research Group. Of particular significance to our future is the formation of the Long Range Planning Committee. The Committee has been very active this year assessing the future of biomolecular resource facilities in terms of potential new directions based on novel, developing technologies and cooperative, systems biology efforts that are beginning to be seen in some institutions. The Committee reported to both the research group chairs and the general membership at the annual meeting in Savannah. Now, based on the feedback of these groups, the Long Range Planning Committee is working on developing milestones for implementing the concepts proposed by the groups.

I also wish to welcome two new Executive Board Members; **Dr. Mike Doyle** and **Dr. Nancy Denslow** who began their terms in December, replacing **Dr. Preston Hensley** and **Dr. Ted Thannhauser.** I believe the association is very fortunate in having Mike and Nancy on the board and I certainly look forward to working with them.

Finally, I invite the membership to feel free to contact me or any other board member if they have ideas or thoughts regarding for the future of the association. It is our wish to ensure that the voice of the members be heard and that the lines of communication are always open.

Sincerely,
Jay W. Fox

ABRF 2006 IN LONG BEACH, CALIFORNIA

Preparation for the 2006 annual ABRF meeting is well underway. The final dates have been set and the meeting will be held February 11–14 in Long Beach, California. **Al Smith** and **Phil Andrews** are the organizers and are busy lining up speakers and sessions. It is not too late to make suggestions or bring forward ideas to either the meeting organizers or Executive Board member **Bill Lane,** as he is the Executive Board liaison for the 2006 meeting organizers.

UPCOMING EVENTS

ABRF members and Corporate Sponsors are encouraged to make announcements about pertinent meetings, workshops, etc. that are beneficial to our members and open to the public. Send items to be listed to Daniel J. Strydom, BioNebraska, Inc., 3820 NW 46th St., Lincoln, NE 68524-1637 (Tel: 402-470-2100; Fax: 402-470-2345; Email: strydom@inetnebr.com).

DATES TO REMEMBER

2005

November 2–4
CLINICAL PROTEOMICS AND BIOMARKER DISCOVERY
Location: North Shore Hyatt, Lake Tahoe, NV, USA
Web Site:
HTTP://WWW.BIOTRAC.COM/PAGES/ONTHEROAD.HTML
Contact Information:
FAES/NIH
Bldg. 60, Rm. 237
One Cloister Court
Bethesda, MD 20814-1460
Phone: 301-496-8290
Fax: 301-402-6292
Email: Mark Nardone; nardonem@mail.nih.gov

2006

February 11–14
ABRF 2006
Location: Long Beach Convention Center, Long Beach, CA, USA
Web Site:
HTTP://WWW.FASEB.ORG/MEETINGS/DEFAULT.HTM

August 27–September 1
17TH INTERNATIONAL MASS SPECTROMETRY CONFERENCE
Location: Prague, Czech Republic
Web Site: HTTP://WWW.IMSC2006.ORG/
Contact Information:
17th IMSC
Conference Secretariat
Institute of Microbiology
Videnska 1083
CZ-142 20 Prague
Czech Republic
E-mail: info@imsc2006.org